70 YEARS OF CREATING TOMORROW

Los Alamos NATIONAL LABORATORY

Mathematical Modeling and Discretization for Exascale Simulation

Luis Chacon,

for the Exascale Mathematics Working Group

07/11/2014

UNCLASSIFIED

Agenda

- Challenges at the exascale
- Exascale computing is all about hierarchies
- Hierarchical models: scale-bridging and coarse graining
- Model coupling and partitioning: accuracy, stability and consistency, solver strategies
- Vlasov-Maxwell: an example
- Other considerations: parallel in time, adaptivity, high-order discretizations

Challenges at the exascale

- While specific architectures are not yet known, exascale computing will bring constraints:
  - Power consumption: will limit data motion and memory access
  - Concurrency: will favor compute-intensive, data-local algorithms
  - Memory: severe bandwidth limitations, better to recompute than to read from memory
  - Data locality
  - Resiliency: soft and hard faults will be common

Exascale machines will be hierarchical

- There will be many levels of parallelism
- Each level of parallelism will bring its constraints, and will require a targeted optimization strategy
  - Each level of parallelism will be best suited for different algorithmic solutions
- The hierarchical nature of the architecture will benefit immensely from hierarchical algorithmic descriptions: multiscale mathematical models

Hierarchical algorithms for hierarchical machines: scale-bridging and coarse graining

- Many applications of interest to DOE are multiscale in nature
  - Tyranny of scales: many orders of magnitude separation between temporal and spatial scales
- Exascale computing can exploit the tyranny of scales to their advantage, provided a suitable algorithmic solution is available: hierarchical algorithms
- Scale-bridging models and coarse-graining strategies will play a key role at exascale.
- Cutting corners for expediency is not acceptable: models must respect physics.

Scale-bridging algorithms

- One can exploit separation of scales to define a model hierarchy (e.g., via coarse-graining): model partitioning
- Model hierarchies based on separation of scales are a good match for exascale computing:
  - Reliable, less intensive levels of description (macroscopic) are mostly unconstrained
  - Most intensive levels of description (microscopic) dominate cost, and will require careful orchestration at most compute-intensive levels
- Model partitioning is not in conflict with tight nonlinear coupling (e.g., nonlinear enslavement)

Scale-bridging algorithms (II)

- Coarse-graining is a natural way to define a model hierarchy:
  - Moment based
  - Homogeneization
  - Renormalization groups
  - Mori-Zwanzig (stochastic PDEs)
- Different levels of the hierarchy may require different discretization approaches, e.g.:
  - Continuum for coarse-grained descriptions
  - Particle-based for fine-grained ones (data parallelism, locality, resiliency, operational intensity)

Model coupling and partitioning

- Strength of coupling among hierarchy levels depends on time scale of interest:
  - Resolving fast time scales will produce non-stiff, weakly coupled systems
  - Stepping over fast time scales will lead to stiff, strongly coupled ones
- When integrating a model hierarchy, one must consider:
  - Solution strategy (loose vs. tight coupling)
  - Propagation of numerical errors across hierarchy (asymptotic well posedness, preservation of conservation laws, nonlinear stability)

Model coupling and partitioning: Partitioned algorithms

- Partitioning can be geometric, operational, and model-based
- Partitioning allows modularity, task parallelism, and reduced synchronization, and is a key element in defining a model hierarchy
- Guiding paradigm: "coupled until proven uncoupled"
- Partitioning enables loose coupling, but is not in conflict with tight coupling (e.g., nonlinear enslavement)

Model coupling and partitioning: Nonlinear solution strategies

- Key for stiff model hierarchies (i.e., for most scale-bridging algorithms)
- Enable strict preservation of conservation laws that depend on coupling across hierarchy levels
- To be practical at the exascale, tight coupling strategies will have to be:
  1. Effectively partitioned (e.g., micro, macro)
  2. Less compute-intensive level drives nonlinear residual (most compute-intensive enslaved), to minimize nonlinear solver memory footprint
  3. Effectively preconditioned (e.g., based on less compute-intensive level)

Model coupling and partitioning: Stability and consistency

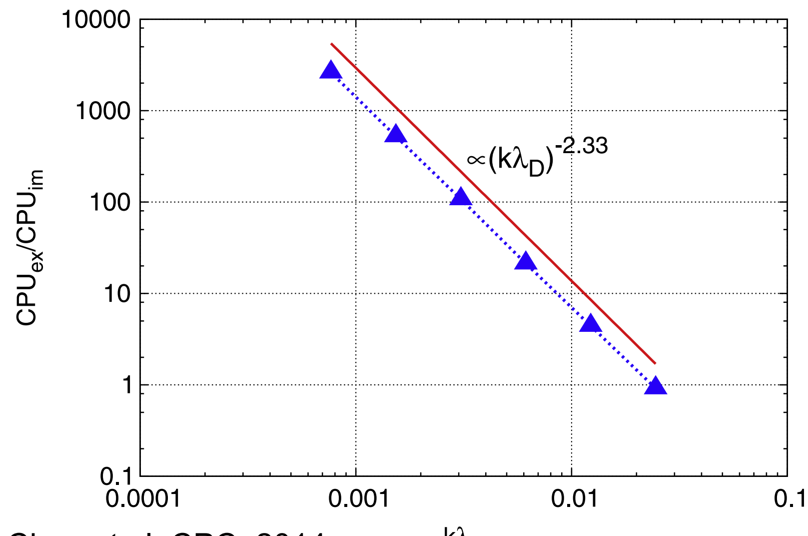
- As critical as ever, if not more!
- Beyond linear stability: consider nonlinear stability
  - Error propagation across levels
  - Preservation of conservation laws (contrains)
  - Asymptotic well-posedness at each level (AP)
- Consistency:
  - High-order is preferred
  - AP property critical (typically low order; needs research)
- Preservation of conservation laws provide many benefits:
  - Local (e.g., soft-faults) vs. global (e.g., nonlinear stab.)
- Particle and stochastic systems present most open questions

Illustrating model coupling and partitioning: Vlasov-Maxwell

- Model hierarchy (moment coarsening):
  - Coarse-grained: Maxwell + fluid moments
  - Fine-grained: Vlasov equation for multiple species
- Model partitioning (tight coupling):
  - Macro fluid system drives nonlinear residual
  - Micro kinetic description is nonlinearly enslaved (auxiliary computation)
- Model discretization:
  - Fluid-field system employs Eulerian representation
  - Kinetic description employs particles
- Hierarchical implementation (co-design):
  - Fluid system implemented on reliable layers (CPU)
  - Particle orbit integration performed on accelerators (GPU)

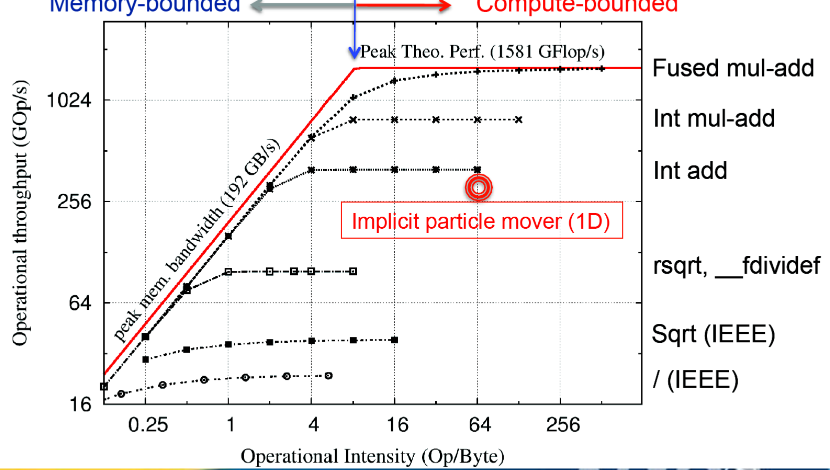
Vlasov-Maxwell: Algorithmic benefits

- Algorithmic state of the art is explicit algorithms, in lock step (memory bounded)
  - Draconian stability constraints, both in time step and mesh resolution!
- Implicit, tightly coupled solve implemented via nonlinear enslavement, driven by fluid/field residual
- Orders of magnitude algorithmic speedup (10^3 demonstrated in 1D, $>10^6$ expected in multi-D)
- Careful co-design cycle renders in in compute-bounded mode, 30% of peak efficiency



Chen et al, CPC, 2014

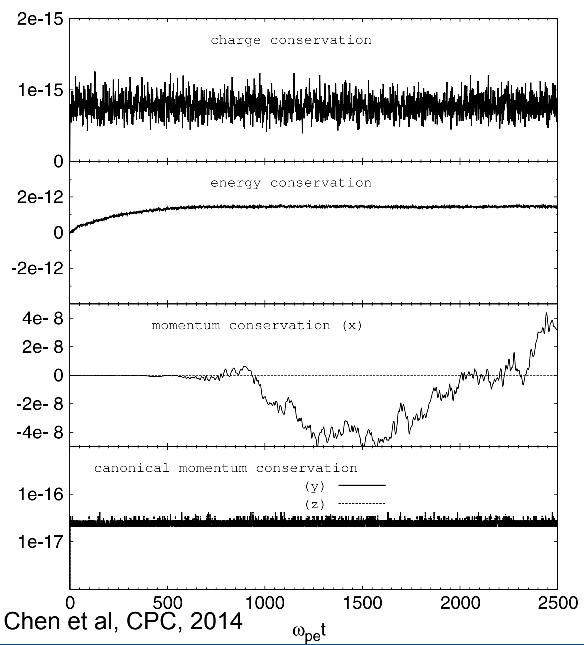
Chen et al, JPC, 2012



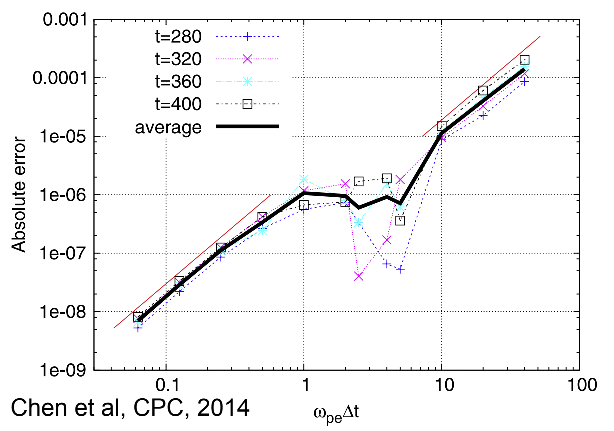
UNCLASSIFIED

Vlasov-Maxwell: Accuracy and stability benefits

- Nonlinear tight coupling enables:
  - Absolute stability
  - Exact preservation of invariants (charge, energy, canonical momenta; a first)
  - Second-order accuracy, with error dominated by slow components of solution (asymptotic preserving)



Chen et al, CPC, 2014



Chen et al, CPC, 2014

UNCLASSIFIED

Unleashing the temporal axis: parallel-in-time approaches

- Sequential aspect of time integration presents a bottleneck
- Need to “open” temporal dimension to iterative treatment
- Many flavors have been explored: parareal (2-level MG), MG-like, SDC-based, etc

Role of high-order discretizations and adaptivity at the exascale

- High-order discretizations promote data locality and operational intensity, and are therefore better suited for the exascale
- Adaptivity will keep playing a fundamental role, but with extended "features"
  - Mesh adaptivity
  - Order adaptivity
  - Model adaptivity
  - Coupling adaptivity

Conclusions

- Exascale computing brings many challenges, but also many opportunities for mathematical exploration
- Algorithms and discrete representations will need to adapt to use these machines
  – Hierarchical architectures will demand hierarchical model descriptions
  – Scale-bridging applications present a significant opportunity
- Exascale computing opens many applied mathematics research questions related to stability, accuracy, asymptotic preservation across levels, and solver strategies